Probatio: A Recommendation System to Assist Educators in Assignment Preparation

Raul Gabrich M. de Freitas, Carla Delgado, João Carlos P. da Silva, Jônatas C. Barbosa and Jean S. Felix
Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

Keywords: Recommendation Systems, Computer Supported Education, e-Learning.

Abstract: Recommendation systems have been used to assist the decision-making processes in a wide variety of fields, such as entertainment, e-commerce and web search engines. Whereas few works have made efforts to assist educators in the elaboration of course assignments, the web system presented in this paper aims at integrating educational metadata and recommendation techniques to support this task. Besides reducing the time required to prepare an assignment, the system can improve the educators' perception of the educational objectives behind it.

1 INTRODUCTION

School evaluation is one of the most important and relevant concepts in the teaching-learning process. From evaluation results, educators can plan, adapt and redesign their teaching activities (Libâneo, 2017). In general, evaluation in educational situations is treated merely as the act of preparing evidence, measuring and making notes, when in fact, it is a broader and more complex process. Under the view of (Stiggins, 2002), the evaluation preparation process should be a continuous evolution of collecting feedback from student results, in order to improve both the learning process as much as future assignments.

To have quality assignments, (Stiggins, 2002) says that: (1) student achievement goals (also known as learning goals, educational objectives and educational goals) must be understood and communicated in advance, (2) the teacher must become instructed in evaluation and assignment preparation, being able to propose exercises and questions that accurately reflect the student's performance, (3) assignments must be used also to build students' confidence in themselves and help them take responsibility for their own learning, in order to establish a basis for lifelong learning, (4) the results of assessments should constitute frequent descriptive feedback for students (rather than critical feedback), (5) instructions should be continually adjusted based on the results of classroom assignments, (6) students should be involved in regular self-assessment, so that they can observe their progress over time and feel in charge of their own success and (7) there should be active communication between students and teachers about their development, status and improvement. In short, the effect of evaluation on learning, as in the classroom, is that students do not give up in frustration or hopelessness, but continue to learn and remain confident that they can reach productive levels.

Thinking of helping educators to produce better assessments, the Probatio system was created. Probatio is a system that aims to assist educators in the management and use of a question bank (assignment items database) and in the preparation of assignments, using artificial intelligence.

The objectives of the Probatio system are:

1. Assist educators to build assignments that are more effective (of better quality) in a more efficient way. This is a short-term goal, as can be achieved since the first use of the system;
2. Improve educators' understanding of the relationship between educational objectives and the quality of assignments, gradually and non-invasive, trying to avoid or minimize resistance. We believe that this is a medium-term goal, as it might be achieve with continuous usage of the system;
3. Offer subsidies and influence teachers to promote re-orientations aimed at improving students' performance in their courses. This is considered a long-term goal, as educators must become confident in the system before accepting it to influence

M. de Freitas, R., Delgado, C., P. da Silva, J., Barbosa, J. and Felix, J.
Probatio: A Recommendation System to Assist Educators in Assignment Preparation.
DOI: 10.5220/0010494203680375
In *Proceedings of the 13th International Conference on Computer Supported Education (CSEDU 2021) - Volume 1*, pages 368-375
ISBN: 978-989-758-502-9

their practices.

We consider effective the assignments whose results (grades obtained) can accurately indicate students' knowledge acquisition level or expected skill development at a certain stage of the course. Quality assessments must be, according to educational objectives, well sized to the expected level of the course, the time available, age and maturity of students – just to mention some of the aspects to be considered. The elaboration of assignments can be facilitated by an interface where the educator easily sees information that pedagogically characterizes an assignment item, such as domain knowledge and cognitive level. This information, as well as information regarding the nature of the item (like time expected to find the correct answer, if it is multiple choice or direct answer) is what we call pedagogical metadata. Though metadata is usually not present in question banks, some can be automatically or semi automatically extracted from the statement of the item or from other metadata.

The selection of questions to compose an assignment can be greatly facilitated and streamlined by a recommendation system. The idea is that, when using the system, the educator is exposed to a series of pedagogical concepts, making her or him reflect on them and, over time, come to a more conscious and intentional use of these concepts during the process of elaborating assignments.

The contributions of this work are: the gathering of information made with educators about the process of preparing assignments and their willingness to use an information system to assist them in this process; and the presentation of the Probatio system, a recommendation system to support assignment preparation. In its current stage, the Probatio system has manual and semi-automatic means to provide metadata for questions. The recommendation engine uses all types of metadata available and is prepared to take advantage of other types of metadata to be include in the future.

Regarding educators' awareness of assignment quality, we consider that the current version of Probatio already brings advances. Currently, Probatio presents the criteria (pedagogical metadata) that can influence the recommendations, so that the user selects the ones to be used. Our results indicate that even this initial stage is already an advance in the routine process of preparing assignments.

The rest of this work is organized as follows: section 2 presents ground work on the subjects of preparing assignments and recommendation systems; in section 3 related works are discussed; in section 4 the Probatio system is presented: interface, architecture and main functionalities; section 5 reports the data obtained after a survey conducted with educators from all levels of formal education in Brazil. Results and conclusion are presented in section 6.

2 GROUND KNOWLEDGE

2.1 Educational Objectives and Their Impact on the Development of Assignments

One of the initial stages of an educational action is to establish educational objectives (Okoye et al., 2013). Educational goals are guidelines that define the expected goal of a curriculum, course or activity in terms of **skills**, **attitudes** or **knowledge**[1] that will be acquired by a student as a result of the process (Krathwohl and Anderson, 2009; Okoye et al., 2013; Simpson, 1966). One way of presenting these skills, knowledge and attitudes is through the elaboration of phrases that have action verbs characterizing the performance or behavior expected by the student in a specific area. *Bloom's taxonomy* is a widely used reference for choosing these verbs and actions (Krathwohl, 2002).

In an introductory programming course, for example, we could establish the following educational goals: (1) **understand** the operation and output produced by short programs that use only the basic programming structures and simple data structures. (2) **create** simple programs from detailed specifications that produce correct results for low complexity problems belonging to the student's universe of knowledge. Thus, from the establishment and dissemination of educational objectives, the teaching-learning process is outlined, including the preparation of assignments (Ferraz et al., 2010).

2.2 Evaluation in the Teaching-learning Process

Although the evaluation process can be seen superficially as the creation and application of assignments (exams), this activity has a much deeper meaning and importance. An evaluation process (assessment) can be defined as an analysis of relevant data in the teaching-learning process that helps educators to make decisions about their work (Luckesi, 2014).

According to (Libâneo, 2017), an evaluation consists of three stages: verification, qualification and

[1]Experts generally use the acronym SKA (*Skills, Knowledge e Attitudes*) to establish educational goals.

qualitative assessment. The verification step consists of collecting data on student achievement through tests, tasks, exercises and other assignments. The qualification stage is the proof of the result obtained in relation to the educational objectives proposed by the educator (when grades are also awarded). Finally, in the qualitative assessment stage, there is a reflection of the work carried out together with the verification of progress and the difficulties encountered in the teaching-learning process as a whole. It is usually at this stage that the actions that lead to didactic replanning are taken. When we take the view of didactic planning as a strategic act, the Probatio system can be considered a decision support system. Probatio assists the teacher directly in the initial two stages of the evaluation process and influences the results that must be analyzed in the third stage.

2.3 Recommendation Systems

Recommendation systems have become a powerful tool to mitigate the problem of information overload (Yang et al., 2003) and boost *e-commerce* sales (Ricci et al., 2015), assisting users in the process of decision making. In addition to recommending assignment questions, which we will discuss in the following section, recommendation systems have already been used in the context of teaching-learning. (Tan et al., 2008) presents a platform for recommending online courses for users looking for teaching materials suited to their needs and interests. (Vialardi et al., 2009) proposes a system for recommending itineraries so that students can choose properly which courses they should enroll based on their past experiences, such as the level of performance in subjects of a certain type or the performance when the weekly workload of courses taken reaches a certain level.

According to (Ricci et al., 2011), the main techniques used for recommendation are content-based recommendation, collaborative filtering and knowledge-based recommendation. In the first, the system is based on items similar to other items that the user has been interested in in the past. In the case of recommending assignment items (questions, activities or problems), such a system could be based on items similar to those used by the teacher in past exams, for example. In collaborative filtering, the system uses information from users with similar interests to recommend items that these users liked / used in the past. In our context, the system could search for questions used by teachers of similar disciplines. Finally, in the knowledge-based recommendation, the system recommends based on the characteristics of the item to be recommended that meet the needs or preferences of the user. Examples are: difficulty level of the question and time needed to answer the question.

3 RELATED WORKS

Cadmus (AIMEUR, 2005) is a hybrid recommendation system to recommend exam questions that uses knowledge-based and content-based recommendation techniques in addition to collecting implicit and explicit feedback from the user to improve their recommendations. Cadmus uses the hybrid recommendation technique (Boulis and Ostendorf, 2005), with a architecture composed of two levels: first level consists of a content-based filter and second level consists of a knowledge-based filter. The content-based system will reduce the search for questions with content relevant to the educator's needs, and the knowledge-based system will order these questions according to the educator's preferences. We consider that the Cadmus assignment preparation process is tiring and repetitive. This is because in addition to specifying the search criteria for the questions, a weight must also be defined for each of the criteria used. Thus, the user needs to fill in more than eight fields, including the definition of weights, to get a recommendation.

Platform PARES (Kaburlasos et al., 2004) was created to deal with the assessment of students in higher education and the absence of continuous assessment throughout the semester. Its goal is making the learning and assessment process a continuous and consistent interaction throughout course time, avoiding that evaluations are concentrated at the end. PARES also proposes to prevent students from plagiarizing their results, generating a set of questions of the same level of difficulty but different in terms of content or ordering in the assessment. However different from Cadmus and Probatio, PARES does not function as a system for recommending questions or using any mechanisms of artificial intelligence or information retrieval to classify the stored questions. It only acts as a facilitator to the teacher in the elaboration of tests, providing an adequate space and model for its creation, and to the students, as a platform used to carry out these evaluations. It is worth noting that PARES was a system developed in mid-2004, about a year before Cadmus.

Several articles (Liu et al., 2018) (Ramesh and Sasikumar, 2010) (Pelánek,) deal with similarity between issues, but this is not the focus of our work. Other works (Jayakodi et al., 2015) (Sangodiah et al., 2016) (Bindra et al., 2017) directly address the issue of automatic question classification, whether in relation to the level of difficulty, content addressed in the

questions or other criteria. These surveys contribute to the study of meta-data on assignment items that can be used as criteria for recommending questions. However, these works in isolation do not address the problem of recommendation or even their association with educational objectives, which is the focus of our work.

4 THE PROBATIO SYSTEM

Probatio is a system that aims to assist teachers in the management and use of a question bank (assignment items' database) as well as in the process of preparing assignments. In short term, Probatio meets the objectives of (1) helping teachers to build better quality assignments in less time, what means a (2) more efficient assignment construction process. In the medium and long term, we have more ambitious and subjective objectives for Probatio, such as: (3) improving educators' understanding of the relationship between the educational objectives of the courses and the quality of assignments; and (4) encourage the continued use of the system so that the educators are able to develop better assessments based on past results. This will contribute not only to the production of better assignments, but also, to better align educators' expectations regarding evaluation scores with the skills developed by their students throughout the course.

4.1 Usability

The main use case for Probatio is the *assignment assembling workflow*. This workflow is supported by the use of recommendation techniques to select questions based on criteria established by educators. Probatio's interface provides groups of criteria of various levels of abstraction - from the most objective, such as the expected time to solve a question - to more sophisticated ones, such as the competence or skill involved in the item. Observing the criteria, educators face different aspects regarding assignment items (what in the least case, increases awareness).

Technically, each instance of these criteria are considered metadata, and mapped to a tag in Probatio. Examples of such tags can be seen in table 1. Note that this are just examples, more tags actually instantiate each criteria. Also, not all criteria apply to all items. What happened is that an item is tagged just with the tags that are applicable, with no compromise to cover all criteria. This flexibility is interesting as the perception an educator has regarding the applicability of a criteria to an item might not come at the time an item is created, but after it has been used a couple of times.

A flexible representation makes it easy to add new metadata to the system. Also, there is no requirement that a criteria is hierarchically superior to another, or that two tags in the same category are mutually exclusive. A lecturer can consider the following criteria to request recommendations: "5 to 10 minutes", "Computer programming", "Nested Loops", "create", "medium". Conceptually, the metadata chosen are possible instances of the following criteria: "time", "knowledge area", "knowledge", "cognitive skill level" and "level of difficulty", respectively (Table 1). The cognitive level tags are related to educational objectives (taken from Bloom's taxonomy as explained in section 2.1).

Besides a selection of tags, an educator might also provide keywords of his or her free choice, which will be searched in the text of the items. The Probatio interface displays, at each stage, the tags and keywords that are being used to generate recommendations. The user has a clear view of the criteria being used to set up the current assignment.

The system also has a *item creation workflow*. This workflow exposes the user to pedagogical information applicable in the process of creating a new item, which will be used later by the recommendation engine in the assignment assembling workflow. This workflow is used to insert new items in the item bank. In figures 1 and 2 we can see an example of a question being created and the association of its meta-information. The item insertion use case foresees that not only freshly created items will be inserted, but also items that are available in other repositories. At insertion time, data like authorship, type (multiple choice or direct answer) and text of the item must be provided. The user can also select related tags (or even create a new tag) and decide whether this item is to be shared with other system users or should be kept private to the creator. All information provided can be changed or complemented later. There is an alternative variation of this workflow: the *item version creation workflow*. This was included in the system as educators said they often create items that are versions or variations of previous items. A relation among versioned items is stored in the item bank.

4.2 System Architecture

The Probatio architecture was based on the Cadmus architecture (AIMEUR, 2005), where the recommendation system has two levels. The two levels interact using the feature augmentation technique, in the same way as Cadmus. Unlike Cadmus, however, the Probatio user profile is extremely simple. It consists of stor-

Table 1: Examples of tags that instantiate some criteria group in Probatio.

criteria (description)	**tags**
time (expected time to answer the item)	1 minute or less; 1 to 3 minutes; 5 to 10 minutes;
knowledge area (as perceived at the institution)	Computer programming; Calculus;
knowledge (formal knowledge involved in the item)	Nested Loops; modularization ; linear function derivative
cognitive skill level (process used to solve the item)	Remember; understand; apply; analyze; evaluate; create
level of difficulty (perceived from past use of the item)	easy; medium; difficult;

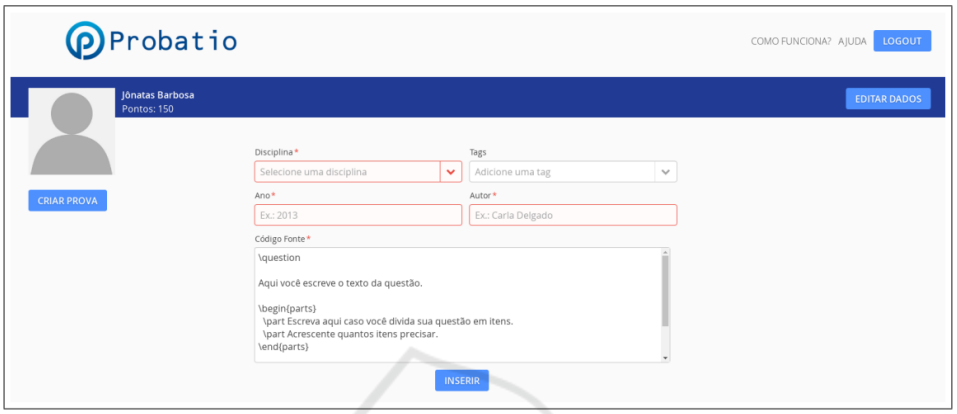

Figure 1: Probation system screenshot: Question creation interface.

ing all the assignments and items already used by the user. Given a user, it is possible to obtain: what assignment has already been issued by him; what items he has created; what items were used in his assignments; and what criteria were previously chosen by that user in the process of recommending items for assembling assignments. Probatio's architecture is outlined in figure 3.

Probatio's first level of recommendation is a **content-based filter** that is responsible for generating a set of candidate questions for the assignment being assembled. In this filter it is possible to select items based on a selection of tags and keywords. The second level of recommendation is a **knowledge-based filter**, responsible for receiving the candidate items generated by the content-based filter and ordering these items according to their relevance to the user. This relevance will be calculated from the comparison between the set of tags for each item and the set of tags that are part of the user profile.

Consider that the users of Probatio are lecturers, teachers, teaching assistants, reviewers, course coordinators or anyone involved in the creation of questions or exams. From a given user, it is possible to retrieve how many times he used a particular tag when preparing his assignments. The more a tag was used by a particular user, the greater the relevance of that tag to him. This information is used to choose the tags that make up the user's profile. Viewing a user's profile allows you to highlight the criteria he or she uses to set up assignments.

5 EDUCATORS PERCEPTION: A SURVEY

In order to better understand the research hypotheses that guided the conception of the Probatio system, we developed a questionnaire whose target audience were educators. Our intentions were: to collect information related to the use of educational objectives during the process of preparing assignments; access the point of view of educators regarding present and future features of Probatio; and validate if there would indeed be an interest in using the system. A questionnaire was prepared with 25 questions, some multiple choice and others discursive. The survey counted with the contribution of 29 educators, 80% of them lecturers of undergraduate courses.

90% of educators stated that written exams were the main resource used in the evaluation of their students. During the process of preparing an assignment item, 72% of educators create and reuse their own items, 69% use items found on the internet, 55% search for textbook items and only 13.8% accept suggestions from colleagues. These first results already show the importance of using written exams in the evaluation process used by teachers, as well as pointing out a possible need for a digital platform for better

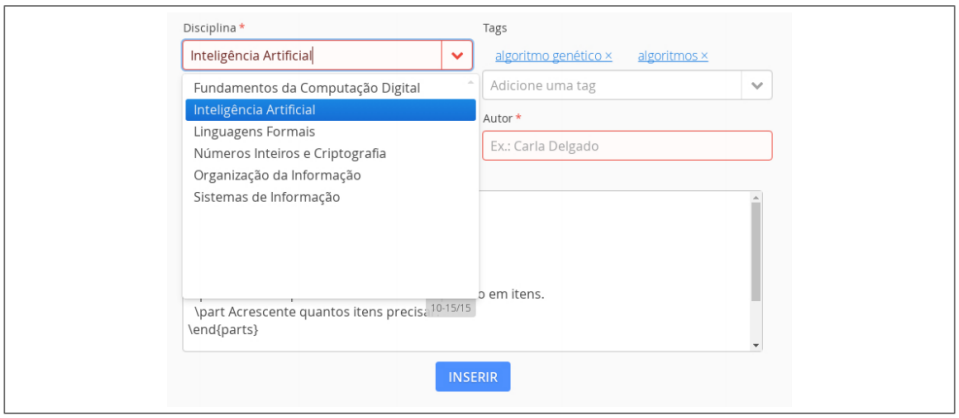

Figure 2: Probation system screenshot: Association of tags (metadata) to a question.

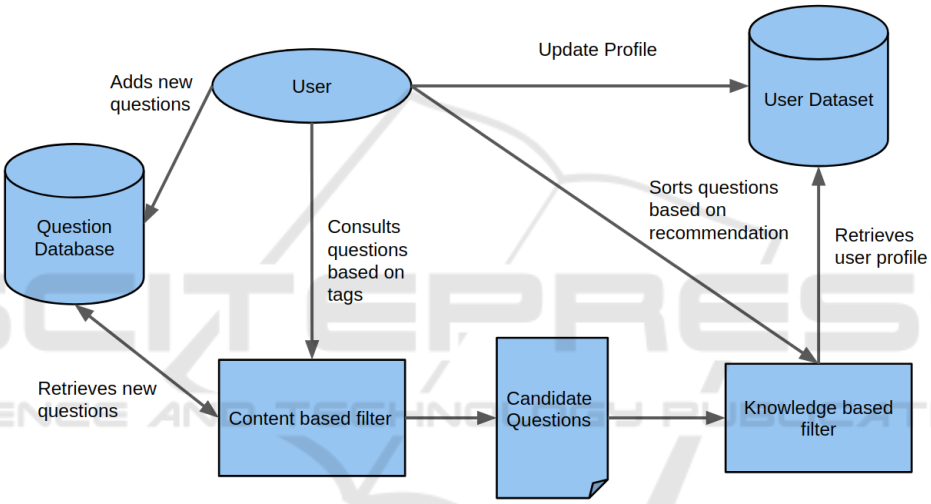

Figure 3: Probatio's system architecture.

storing and managing their items. Thus, when asking whether these educators use any digital medium to store their old exams, 96.4% of respondents said yes. 72.4% think it is a good idea to have a specific web platform capable of storing their assignments (exams, exercise sheets,..) and previously created/used items to be used as a reference in future assignments.

Regarding the establishment of educational objectives for an assignment, 51.7% of the educators stated that they always set these objectives, 35.4% marked that most of the time yes and the rest never or almost never. After elaborating the assignment, 59% states that he or she never had his assignment reviewed by anyone before handing it over to students. Many find the assignment assembling process time consuming (60.7%), but important (71.4%). It is interesting to note that, in addition to the majority agreeing that the assignment preparation process takes time, practically everyone believes that this process does not get faster with time (96.4%). Although many educators perceive the time invested and the importance of preparing a good assignment, only 42.9% see a direct relationship between the time spent on preparing the assignment and the results obtained by the students.

At the end of the questionnaire, we elaborated some questions aimed exclusively at evaluating Probatio's functionalities. With regard to question sharing, 58.6% say they often share their questions with other teachers and 89.3% likes or finds it useful to have other teachers share their questions. Regarding the recommendation of these questions stored in a bank, 65.5% would like to receive recommendations for questions, 27.6% marked this option as "perhaps" and only 6.9% would not.

Finally, we left a space for suggestions or restrictions that the interviewees considered important for the platform to have. Some suggestions such as question authorship, validation of the teacher's identity to

separate him from the students, a friendly interface and an efficient search filter to search for questions are among the most relevant answers. Thus, this research was of great value for a brief validation of the current state of the platform and targeting future features.

The data collected endorse the hypotheses that guided the conception of Probatio. With the use of Probatio, the time used in preparing a test can be drastically reduced. This is due to the ease of handling and retrieving questions, the reliability of the stored questions, due to the revisions that the questions could undergo by educators, in addition to the use of the recommendation system added to the platform. In addition to the main recommendation feature, other features provided for in Probatio also appear to be on the teachers' "wish list". The questions stored in the bank could be reviewed and evaluated by the teachers who use it, thus increasing the reliability in using that particular question.

6 CONCLUSIONS

We see Probatio as an innovative tool. In its current state, Probatio allows questions to be stored in its question bank and retrieved both by a search process and by recommendation. Tags and keywords are used to recommend items through a content-based filter and later, as a knowledge-based recommender. Users can retrieve items to prepare their assignments using a simple interface where they state "what they want" (criteria for the recommendations) and select, amount the retrieved items, the ones he or she wants in the assignment.

Our recommendation system is already able to deal with several metadata, but by now, metadata are manually associated to items. As future work, we intend to implement automatic extraction of relevant metadata from the items' text. We are currently evaluating the use of machine learning to categorize questions in the cognitive levels of Bloom's taxonomy. Many works already investigated such categorization of questions written in the English language (Bindra et al., 2017; Sangodiah et al., 2016; Sangodiah et al., 2014). As the system is being used in Brazil, a classifier for Portuguese language is needed. Each new metadata made available in Probatio must be explicitly incorporated in the interface, so that the link between the chosen criteria and educational objectives is always in focus.

Automatic feedback could be used in the future to improve semantic information. For instance, the collection of the percentage of students who correctly answered each question could help to classify the questions as "easy", "medium" or "difficult", or even in more sophisticated categories in case the type of mistake could be automatically identified.

It is also in our agenda the creation of simple dashboards where users of the system can have a glimpse of the tags associated to each item, the most relevant tags in an assignment, and the tags most used by an educator.

It is believed that the application of an information system to support assignment creation and the management of assignment items has a positive effect not only for educators, but also for students and educational institutions. The item bank itself, storing items, educational metadata and relations among them is a valuable asset for the institution. Users of the system are expected to improve their educational skills with time. Our most ambitious goal is for teachers to learn more about educational objectives and criteria for setting up assignments, and to make better use of the feedback that an assessment can provide. In the long run, it is expected that the continued use of the system will provide a maturation of educators' understanding and perception of the results of his evaluations.

REFERENCES

AIMEUR, H. H. E. (2005). Exam question recommender system. *Artificial Intelligence in Education: Supporting Learning Through Intelligent and Socially Informed Technology*, 125:249.

Bindra, S. K., Girdhar, A., and Bamrah, I. S. (2017). Outcome based predictive analysis of automatic question paper using data mining. In *2017 2nd International Conference on Communication and Electronics Systems (ICCES)*, pages 629–634. IEEE.

Boulis, C. and Ostendorf, M. (2005). Text classification by augmenting the bag-of-words representation with redundancy-compensated bigrams. In *Proc. of the International Workshop in Feature Selection in Data Mining*, pages 9–16. Citeseer.

Ferraz, A., Belhot, R. V., et al. (2010). Taxonomia de bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. *Gest. Prod., São Carlos*, 17(2):421–431.

Jayakodi, K., Bandara, M., and Perera, I. (2015). An automatic classifier for exam questions in engineering: A process for bloom's taxonomy. In *2015 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, pages 195–202. IEEE.

Kaburlasos, V. G., Marinagi, C. C., and Tsoukalas, V. T. (2004). Pares: A software tool for computer-based testing and evaluation used in the greek higher education system. In *IEEE International Conference on Advanced Learning Technologies, 2004. Proceedings.*, pages 771–773. IEEE.

Krathwohl, D. R. (2002). A revision of bloom's taxonomy: An overview. *Theory into practice*, 41(4):212–218.

Krathwohl, D. R. and Anderson, L. W. (2009). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Libâneo, J. C. (2017). *Didática*. Cortez Editora.

Liu, Q., Huang, Z., Huang, Z., Liu, C., Chen, E., Su, Y., and Hu, G. (2018). Finding similar exercises in online education systems. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1821–1830. ACM.

Luckesi, C. C. (2014). *Avaliação da aprendizagem escolar: estudos e proposições*. Cortez editora.

Okoye, I., Sumner, T., and Bethard, S. (2013). Automatic extraction of core learning goals and generation of pedagogical sequences through a collection of digital library resources. In *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries*, pages 67–76. ACM.

Pelánek, R. Measuring similarity of educational items: An overview.

Ramesh, R. and Sasikumar, M. (2010). Use of ontology in addressing the issues of question similarity in distributed question bank. In *Proceedings of the International Conference and Workshop on Emerging Trends in Technology*, pages 700–703. ACM.

Ricci, F., Rokach, L., and Shapira, B. (2011). Introduction to recommender systems handbook. In *Recommender systems handbook*, pages 1–35. Springer.

Ricci, F., Rokach, L., Shapira, B., and Kantor, P. B. (2015). *Recommender systems handbook*. Springer.

Sangodiah, A., Ahmad, R., and Ahmad, W. F. W. (2014). A review in feature extraction approach in question classification using support vector machine. In *2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014)*, pages 536–541. IEEE.

Sangodiah, A., Ahmad, R., and Ahmad, W. F. W. (2016). Integration of machine learning approach in item bank test system. In *2016 3rd International Conference on Computer and Information Sciences (ICCOINS)*, pages 164–168. IEEE.

Simpson, E. J. (1966). The classification of educational objectives, psychomotor domain.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10):758–765.

Tan, H., Guo, J., and Li, Y. (2008). E-learning recommendation system. In *2008 International Conference on Computer Science and Software Engineering*, volume 5, pages 430–433. IEEE.

Vialardi, C., Bravo, J., Shafti, L., and Ortigosa, A. (2009). Recommendation in higher education using data mining techniques. *International Working Group on Educational Data Mining*.

Yang, C. C., Chen, H., and Hong, K. (2003). Visualization of large category map for internet browsing. *Decision support systems*, 35(1):89–102.